**A Short Report: Decision Tree Classification on Breast Cancer Dataset**

**Objective:** Classify breast tumors as benign or malignant using the J48 Decision Tree in Weka.

**Introduction:** Breast Cancer is one of the most common cancers affecting women worldwide. Machine learning techniques like decision trees help classify tumors based on medical attributes. This report explores the J48 algorithm on the Wisconsin Breast Cancer Dataset.

**Purpose of Decision Tree:**

A decision tree is a supervised learning algorithm that splits data into branches based on feature values, forming a tree like structure that leads to class labels.

The most important features in diagnosing breast cancer are determined based on their position in the decision tree which includes Cell Size Uniformity,Bare Nuclei,Clump Thickness,Cell Shape Uniformity,Bland Chromatin and Marginal Adhesion.

The key advantages of using decision trees include: Interpretability, Feature Importance, On-Linearity Handling, Minimal Data Preprocessing.

**Model Performance Analysis:**

**-Accuracy and Performance Metrics:**

The trained J48 decision tree achieved an accuracy of 95.14 % meaning it correctly classified 665 out of 699 instances. The detailed performance metrics are as follows:

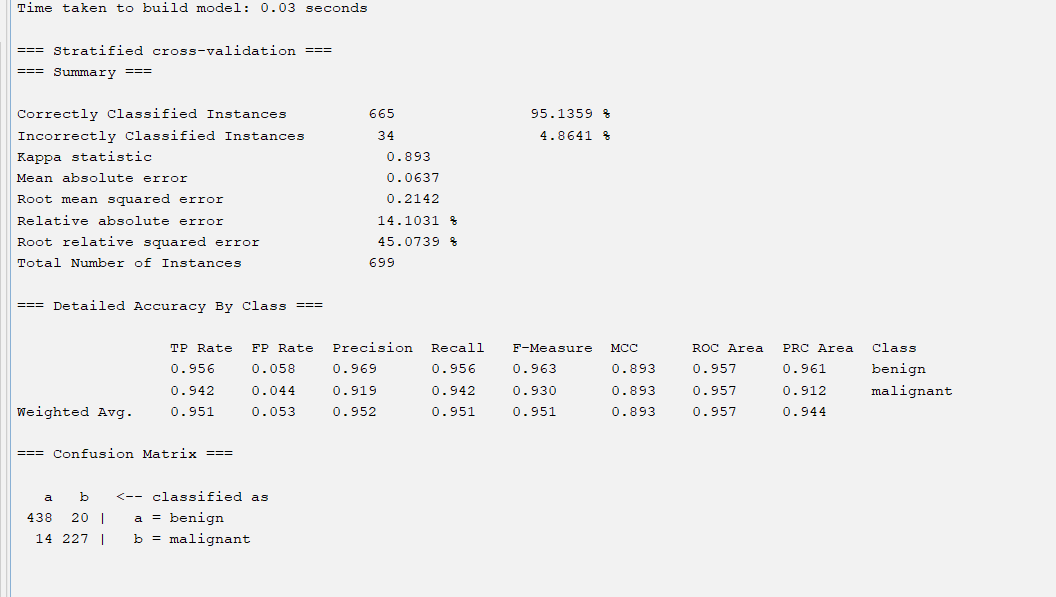
* **Sensitivity (Malignant):** 0.942
* **False Positive Rate:**0.044
* **Precision (Benign: 0.969, Malignant: 0.919)**
* **F1-Score (Benign: 0.963, Malignant: 0.930)**
* **ROC Area**: 0.957

**Confusion Matrix Interpretation:**

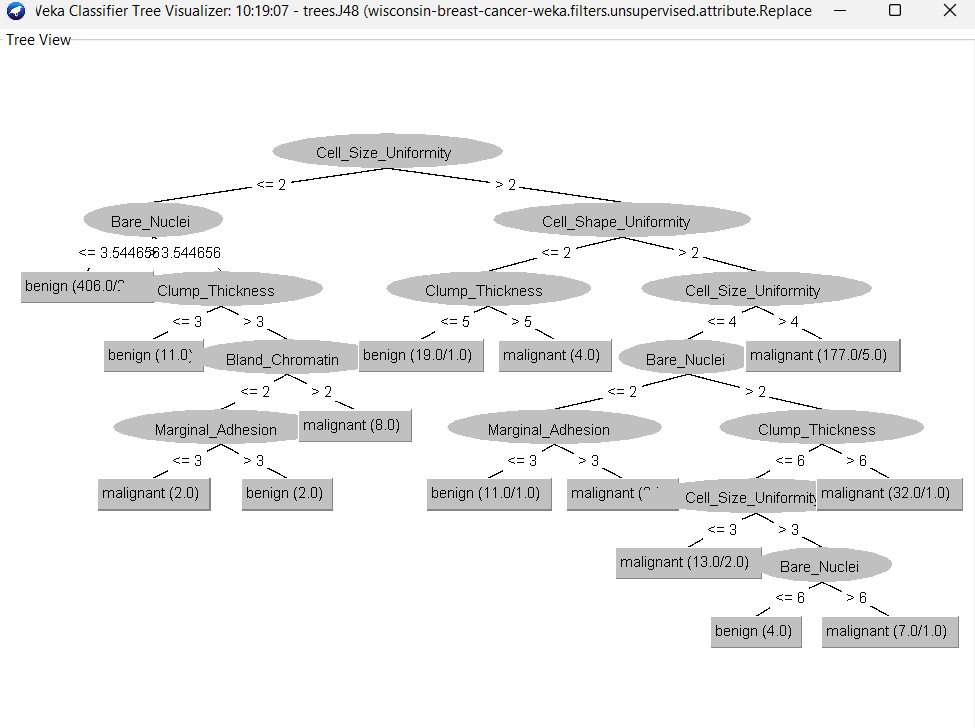
✔ **True Positives (438)** – Correct Benign cases  
✔ **True Negatives (227)** – Correct Malignant cases  
✘ **False Positives (14)** – Malignant misclassified as Benign  
✘ **False Negatives (20)** – Benign misclassified as Malignant

**Screenshots:**

**1.Classifier Output (Accuracy, Confusion Matrix and Evaluation and Metrics)**



**2.Decision Tree Visualization**



**Conclusion:**

The J48 classifier demonstrates **high accuracy (95.14%)**, making it a valuable tool for medical diagnosis. The low false negative rate is crucial in healthcare applications. Decision trees provide an interpretable and effective approach for breast cancer classification, aiding medical professionals in early detection and treatment.

Submitted By : Aaditya Sapkota, BDSc Roll No: 37